

Modeling and Simulation in Biochemistry

A guide for users and consumers
of crystallographic information

Katherine A. Kantardjieff, Ph.D.
Keck Center for Molecular Structure
California State University Fullerton

Outline

- Crystallography education policies
 - American Crystallographic Association
 - United States National Committee for Crystallography
- Remotely-enabled instrumentation
 - CMoIS
 - PRISM
 - NSF
- Biomolecular crystallography 101
- Structure model validation
 - Seven criteria
- Ongoing experiments in the biochemistry laboratory
 - Lactate dehydrogenase complexes
- Take home messages
- Acknowledgements

ACS Washington DC © KAK 2009

<http://aca.hwi.buffalo.edu/List.html>



Molecular Structure by Crystallography

Crystallography Education Policies for
the Physical and Life Sciences

Sustaining the Science of Molecular Structure in
the 21st Century

Prepared by the American Crystallographic Association
and the United States National Committee for Crystallography
©2006

ACS Washington DC © KAK 2009

Recommendations for Undergraduate Students in the Life Sciences

- **Intellectual Context**
 - An undergraduate student will appreciate the extraordinary wealth of information provided by structural biology and the impact of crystallography on life sciences.
- **Basic Concepts of X-rays and Scattering**
 - An undergraduate student will understand the concepts necessary to appreciate the powerful results, as well as the advantages and limitations of the method.
 - An undergraduate will have a qualitative understanding of the interaction of electromagnetic wave with periodic arrays, diffraction and reconstruction to form an electron density image, symmetry in nature and its use in the assembly of macromolecules.
- **Data Usage and Model Evaluation**
 - All competent users of X-ray models will be able to obtain and view them from databases. More advanced users should have detailed knowledge of them.

ACS Washington DC © KAK 2009

Recommendations for Undergraduate Students in the Life Sciences

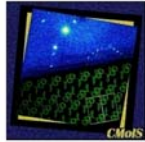

- **Molecular Concepts**
 - Undergraduates should have a solid understanding of the fundamentals of three-dimensional structure in order to communicate and use these concepts in the context of other courses in the curriculum and in their own inquiry and investigation.
 - The function of the macromolecule is determined by the 3D arrangement of the atoms.
 - Biomolecular structures derived from crystallographic analysis can be used to develop students' understanding of protein interaction at the molecular level with other proteins, nucleic acids, small molecules, drugs, and substrates.
- **Advanced Concepts and Applications**
 - Biomolecular structure provides the molecular basis of human health and disease.
 - The results of protein crystallographic analysis may be used to further develop students' understanding of the molecular basis of heredity and biological evolution, as well as matter, energy and organization of living systems.

ACS Washington DC © KAK 2009

CMoIS

(<http://chemsrvr2.fullerton.edu/CMOLS/index.htm>)

- Conceived in 1992, established in 1994
- Third of the CSUPERB core facilities
- First facility of its kind at a predominantly undergraduate institution
- Dedicated to molecular structure determination and analysis using X-ray diffraction and computational approaches
 - Computational facilities online in 1995
 - X-ray instrumentation available in 1996
 - **Instrumentation remotely accessible to the CSU since 1997**
- One of five core facilities of the STaRBURSTT CyberDiffraction Consortium (established 2005)
- Funding
 - State of California, CSUPERB, Camille and Henry Dreyfus Foundation, National Science Foundation, W.M. Keck Foundation
- Corporate sponsorship
 - Active Sight, Bruker AXS, CAChe (Fujitsu), deCODE Biosciences, Emerald Biosystems, GE Healthcare, H. Research, Molecular Dimensions Limited, Molssoft LLC, Nextal (Qiagen), Oxford Diffraction Systems, Rigaku

ACS Washington DC © KAK 2009

6



Partnership for Remote Instruments to Study the Structure of Matter

X-ray diffraction
Inductively coupled plasma mass spectrometry
Scanning electron microscopy
Atomic force microscopy
Confocal microscopy
Nuclear magnetic resonance
Electron paramagnetic resonance
Circular dichroism

CAL STATE FULLERTON
CALIFORNIA STATE UNIVERSITY LONG BEACH
CAL POLY POMONA
FRESNO STATE
California State University, Fresno

California State University Northridge

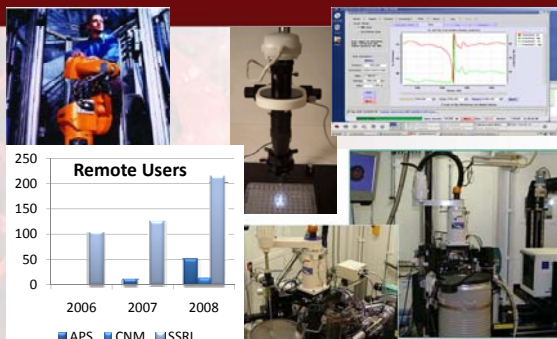
HARVEY MUDD COLLEGE

Newport Harbor High School
A CALIFORNIA DISTINGUISHED SCHOOL

ACS Washington DC © KAK 2009

7

Synchrotrons, robotics, and automation



Remote Users

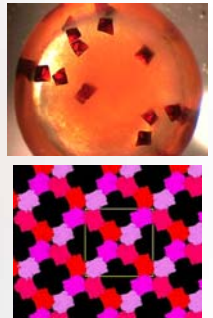
Year	APS	CNM	SSRL
2006	100	20	10
2007	120	30	15
2008	200	40	20

ACS Washington DC © KAK 2009

8

Crystallography myths and realities

- Regarded particularly by the biochemical community as "gospel truth" – database user beware! BUT...
- Dominant method
- It is not monkey work; it is not like running a spectrometer.
- Elastic emission of radiation due to acceleration of electrons in response to an incident electromagnetic wave
- Scattered waves interfere
 - Total scatter from object is exquisitely sensitive to internal structure
- X-ray wavelength is of the order of interatomic dimensions
- Diffraction from crystals concentrates the scattering into discrete directions and amplifies the signal
- Structure obtained is an average of all the molecules in the crystalline lattice (NOT purely a static picture)

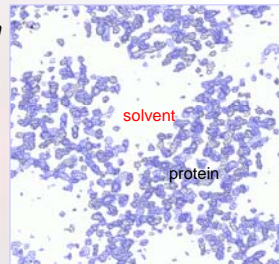


ACS Washington DC © KAK 2009

9

Crystallography myths and realities

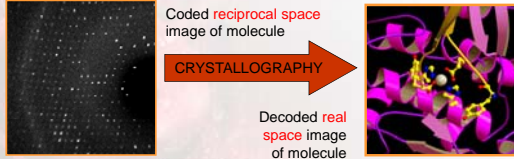
- Virtually everything we know about biochemical mechanism comes from crystallographic information*
 - The lattice is highly aqueous
 - Molecules usually adopt native conformations
 - Differences may arise due to packing contacts (look at polymorphs)
 - Differences may arise due to presence or absence of ligands/cofactors



ACS Washington DC © KAK 2009

10

Crystallography: deciphering the reciprocal space code



Cracking the crystallographic code:

Nature of code: diffraction physics and reciprocal space

Statistical analysis: intensity statistics and distributions, probabilities, outliers

Patterns: symmetries, phase relations, bootstrap

Context: prior knowledge, sequence, physical data, chemical knowledge

ACS Washington DC © KAK 2009

11

Crystal structure determination

- Nature of diffraction:** Reciprocal space and lattice, geometry, convolution and the basic concept of FT
- Interaction of X-rays with matter:** Matter determines the contents of pattern, atomic scattering, structure factors, B-factors, *atomistic view*.
- What information is where in the data**
 - Intensities inform us about types of atoms and nature of the underlying crystal lattice.
 - Last experiment before computation, most important (time = \$)
- Phase problem**
 - We don't have lenses that focus X-rays. We record intensities of scattered waves (amplitude² = I), but we don't know relative phases to constructively and destructively interfere waves to produce an image of the object (molecule) doing the scattering
 - We recreate the action of the lens mathematically on the computer using a variety of methods
 - Molecular replacement
 - Heavy atom location using isomorphous and anomalous techniques - *substructure*
- Model building and refinement**
- Structure validation and analysis**

ACS Washington DC © KAK 2009

Model validation

- Fundamental necessity of structure guided bioinformatics, drug design and computational methods is that structure model of target is available
- Models may be derived experimentally
- Models may be derived *in silico* (homology modeling) based on experimental data
- Quality of underlying model critically determines the validity and accuracy of interpretation or *in silico* results**
- Each *in silico* method poses different requirements with regard to the underlying molecular model**

ACS Washington DC © KAK 2009

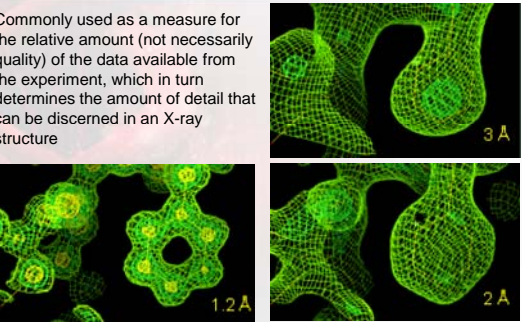
Model validation

- Model generally refers to a representation of a molecular structure, experimentally derived or computational, that agrees with and explains what we know from supporting experiments
- Quality
 - General rule: an X-ray structure is better than a NMR structure, because NMR structures usually emphasize dynamics at the expense of coordinate precision
 - Quality of a computational model will depend on its sequence and fold similarity to the underlying crystal structure template, and the quality of that template
- Necessary to assess the merits of each structure model in relation to the task at hand
 - A computational model based on a highly homologous, well determined X-ray structure may well rival or exceed the accuracy of an experimental NMR structure, and be entirely satisfactory for an *in silico* mutation study or virtual screening

ACS Washington DC © KAK 2009

Criterion #1 - Resolution

- Commonly used as a measure for the relative amount (not necessarily quality) of the data available from the experiment, which in turn determines the amount of detail that can be discerned in an X-ray structure



ACS Washington DC © KAK 2009

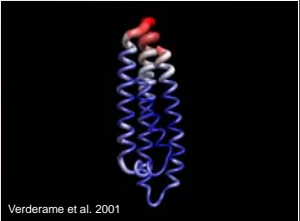
Criterion #2 – R value and free R

- A measure for the fit of the calculated versus the experimental diffraction data
- General rule: Lower R value and smaller gap between R value and free R indicate better overall quality of the model. R values for high resolution models tend to be lower, as the model can be described in more detail and thus tends to represent a more accurate description of reality.
- Global quality indicators: Define how well *all* observed intensities (or structure factor amplitudes) in the experimental diffraction pattern match the structure factor amplitudes calculated from *all* the atoms in the model
- Local definition of quality is highly relevant for accurate biological interpretation around active sites, mutations, and bound ligands (essential for successful structure-guided drug discovery).

ACS Washington DC © KAK 2009

Criterion #3 – B factors

- Debye-Waller temperature factor can be formally interpreted as a measure of the thermal mobility of the corresponding atom via the mean square positional displacement ($B=8\pi^2\langle u^2 \rangle$)
- A large deviation from average B-factors is usually an indication of lower certainty of the positional precision and accuracy in the corresponding region
- B-factor column in computational models is based on the amount of structural information that supports each part of the model (measure of confidence)

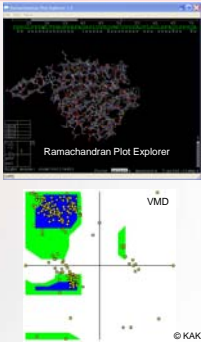


Verderame et al. 2001
PDB ID: 1EA8 rendered with UCSF Chimera

ACS Washington DC © KAK 2009

Criterion #4 – Geometry-based validation

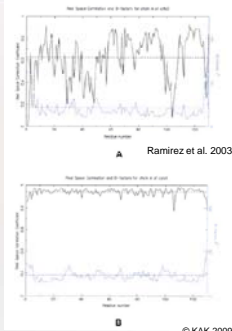
- Measures to what degree a model structure deviates from expected bond lengths and angles
- Backbone torsion angles represented by the Ramachandran plot
 - PDBCHECK, PDBREPORT, PROVE and WHAT_IF



ACS Washington DC © KAK 2009

Criterion #5 – Validation against experimental data ¹⁸

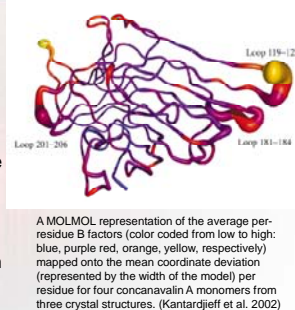
- **Real space correlation coefficient (RSCC) plots** depict, residue by residue, the fit of the model (including ligands) to the electron density. Weak correlation indicates poor fit to electron density, either as a result of genuine absence of ordered regions, or possibly due to improper model building, and thus a lack of reliable information in that part of the structure
- A service that allows this assessment for already deposited structures (provided structure factors were also submitted) exists as the electron density server at the University of Uppsala <http://portray.bmc.uu.se/eds/> and at the EBI/MSD



ACS Washington DC

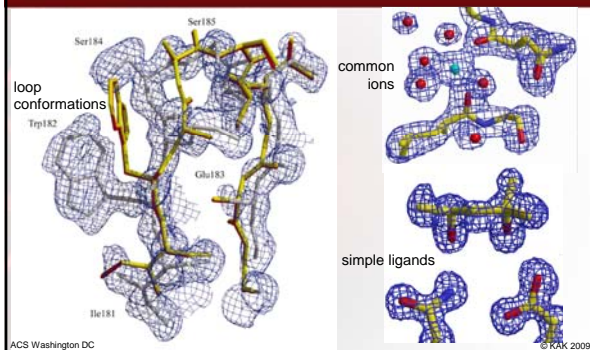
Criterion #6 – Crystal packing and quaternary structure ¹⁹

- In the vicinity of crystal contacts, conformational changes can occur, which may affect the relevance of any biological interpretation, particularly if these regions are of functional significance
- Users of X-ray structure models (PDB files) should display symmetry related molecules to recognize such possible complications



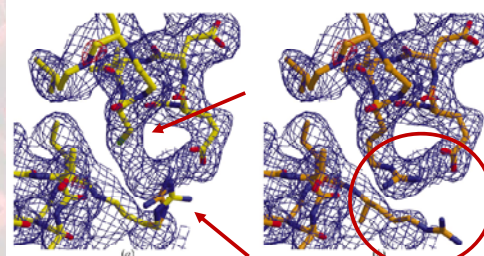
ACS Washington DC

A closer look at some electron density maps (Kantardjieff et al. 2002) ²⁰



ACS Washington DC

A closer look at some electron density maps (Verderame et al. 2001) ²¹



ApoE3/4 isoforms differ in a single mutation C112R

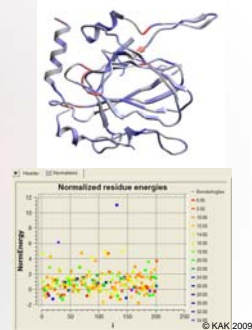
Mutation disrupts ion pair with E109 and extended R61 at helix 2-3 interface.

ACS Washington DC

© KAK 2009

Criterion #7 – Energy and fold potential-based methods ²²

- Particularly valuable for low homology computational models (and also low resolution X-ray and NMR structures)
- Assist in determining whether one has produced a correct structure by indicating potentially misfolded parts of the structure
- ICM, PROSAIL, Verify3D, ANOLEA
- ICM 'protein health' compares the semi-empirical force field interaction energy of a single amino acid residue with the entire protein structure and surrounding solvent with an energy distribution observed in high-resolution crystal structures



ACS Washington DC

Structural biology of LDH – a contemporary biochemistry laboratory ²³

- A contemporary experiment in macromolecular structure/function analysis
 - advanced crystallographic techniques and methods
 - front end aspects of protein crystallography: protein production, purification, and crystallization
 - back end aspects of structure validation and analysis
- Took an existing experiment in the senior biochemistry laboratory at CSU Fullerton and expanded it into a comprehensive structure determination and analysis by x-ray diffraction
- LDH from chicken breast muscle (LDH-A) is purified, biochemically and biophysically characterized, crystallized, and its structure has been determined for the first time by x-ray crystallography
- Exposes students to the modern methods of structural genomics
- Provides new insights into the structural features of this LDH that govern its kinetic and stability properties
- Lactate dehydrogenase (LDH) is a well-studied essential enzyme in carbohydrate metabolism
 - extensive amino acid data are available for orthologous homologs
 - some atomic resolution structure information is available

ACS Washington DC

© KAK 2009

LDH

24

- LDH (EC 1.1.1.27) is an oxidoreductase, which catalyzes reduction of pyruvate to lactate by oxidizing NADH to NAD⁺:
- Students isolate LDH-A from chicken breast muscle using standard techniques
- LDH activity is assayed spectrophotometrically using established protocols
- Students perform homology modeling and in silico mutagenesis using ICM-Pro (Molsoft).
- Students perform an extensive crystallization screening on the native protein and complexes with lactate, pyruvate, NADH and the inhibitor, oxalate.
- Crystals are examined for diffraction quality. Conditions are optimized as needed.
- Crystals are cryoprotected, mounted in cryoloops, placed in cryovials, placed in a shipping dewar and sent to one of three synchrotrons for remote data collection.

ACS Washington DC © KAK 2009

LDH

25

- Data are scaled and reduced locally
- Phasing is performed by molecular replacement
 - Native enzyme structure was phased using coordinates of either the porcine or the human enzyme
 - Complex structures have been or will be determined using the coordinates of the determined native structure
 - Phase bias removal procedures are performed during refinement
 - Model building and refinement performed with the CCP4 suite
- Students are engaged at every level of structure determination

ACS Washington DC © KAK 2009

Results

26

ACS Washington DC © KAK 2009

Take home messages...

27

- Crystallographic models validated by established criteria can provide valuable insights into biological structure and function**
- The quality of the underlying model critically determines the validity and accuracy of mechanistic interpretation or in silico results**
 - The *accuracy* of the model presented and subsequently interpreted by the calculated electron density map depends on the underlying data quality, phasing and refinement by the crystallographer (GIGÓ). This is not to be confused with *precision* and level of detail.
- Models should agree with mutational data, as well as experimentally derived structure-activity relationships**
 - Does the model make biological, chemical and physical sense?
- Computational approaches (structural bioinformatics, homology modeling, virtual screening) can be valuable tools of the X-ray crystallographer and the experimentalist**
 - Caveat emptor

ACS Washington DC © KAK 2009

Acknowledgements

28

- ACA
- USNC/Cr (NAS/NRC)
- Department of Chemistry and Biochemistry
 - Chandra Srinivasan, Chris Meyer, Fu-Ming Tao, Nephi Polder, Go Watanabe, Teddy Kidane
- College of Natural Science and Mathematics
- Information Technology
- CSUPERB
- iLinc™ Communications
- The Boeing Company
- National Science Foundation
- W.M. Keck Foundation
- Larry Grant, Erin Greiner, Laura Ramirez, Joe Verderame
- StARBURSTT-CDC
 - Allen Hunter, Greg Ferrence, Guy Crundwell, Marcus Bond
- PRISSM
 - Alejandro Calderon-Ureda, Tim Corcoran, Vikki Costa, Amy Cox-Petersen, Julie Karjala, Zed Mason, Phyllis Nelson, Shantanu Sharma, Koni Stone
- SSRL
 - Herb Axelrod, Clyde Smith, Michael Soltis, Thomas Weiss
- CMoS staff
 - Guy Crundwell, Brian Schick, Steve Herron, Sean Ouyang, James Zhao

ACS Washington DC © KAK 2009

Questions...

29

- kkantardjief@fullerton.edu
- <http://faculty.fullerton.edu/kkantardjief>
- <http://ilincusers.blogspot.com>
- <http://twitter.com/kkantar>
- <http://www.facebook.com/kkantardjief>
- <http://www.cei2008workshop.org>
- <http://aca.hwi.buffalo.edu/List.html>

ACS Washington DC © KAK 2009